DOCUMENT RESUME

ED 388 110 FL 023 388

AUTHOR Schackne, Steve
TITLE Extensive Reading and Language Acquisition: Is There a Correlation? A Two-Part Study.
PUB DATE Dec 94
NOTE 28p.; Paper presented at the Annual International Conference of the Institute of Language in Education (Hong Kong, December 1994).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Educational Strategies; *English (Second Language); Foreign Countries; Higher Education; Instructional Effectiveness; *Reading; Reading Instruction; *Recreational Reading; Second Language Learning; *Second Languages

ABSTRACT

A 1986 study concerning the effectiveness of extensive reading in improving second language learning, and its replication in 1994-95, are reported. In the original study, four classes of English as a Second Language in a Taiwan university were used as experimental and control groups, the only difference in instruction being the use of extensive reading for pleasure in one. All experimental classes showed greater gains in reading skills. A study using both the same and additional measurement instruments and a much larger sample was undertaken at that university and another in Macau; results at the latter institution are reported here. Again, experimental group gains were greater than control group gains, but at a lower level of significance. Possible explanations for this discrepancy in results are discussed. A 29-item list of student readers is included. Contains 22 references. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 388 110

EXTENSIVE READING AND LANGUAGE ACQUISITION
IS THERE A CORRELATION?

A TWO-PART STUDY

S'eve Schackne
University of Macau

	Type	Locale	Year
PART I	Original Study	Tunghai University Taichung, Taiwan	1986

	Type	Locale	Year
PART II	Replication Study	University of Macau Macau	1994

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Steven Schackne

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL023388

BEST COPY AVAILABLE

Reading for Pleasure and Language Acquisition

Stephen Schackne

University of Macau

The following experiment, conducted in 1986, seeks to quantify an assumption that has long been prevalent in ESL-EFL circles, namely, that extensive reading enhances language acquisition.

The experiment was controlled, the experimental and control group differing only in treatment (extensive reading).

The results showed the experimental group (which read extensively) outdistanced the control group (which read from an intensive reading skills textbook, In Context) along both raw and percentage measures +60%>159%

The author concluded that there was statistical evidence to support a correlation between extensive reading and language acquisition, and touched on the practical considerations of such a correlation.

The article closed by posing a question: if there is both popular and empirical support for extensive reading in language learning, why isn't it both encouraged and used in more EFL-ESL programs throughout the world?

This 1986 article was not published in anticipation of a later opportunity to test the results by replication. I now present the original study with some revision, accompanied by the replication study of the 1994-1995 academic year.

Background

I first became interested in extensive reading in 1984 while in a graduate TESOL program at the State University of New York at Albany. The Chinese students in the program possessed a communicative competence which appeared to surpass that of their counterparts from other countries. Upon questioning, these students yielded remarkably similar answers--their schools used audio-lingual drills and extensive reading as the main foundations of their English programs. Could the audio-lingual drills account for their superior mastery? I doubted it. Many of the other students were trained primarily through the audio-lingual method, and the results were mixed. Could it be the extensive reading? Possibly. None of the other students were trained by this method. Also, it made intuitive sense--my American friends and I all agreed that the really accomplished practitioners of the written and spoken word that we knew were all extensive readers as well.

The literature yielded nothing specific. Most of the articles focused on the teaching of reading and reading strategies. And when I did run into articles on extensive reading (Jiang 1984; Hubbard, Jones, Thornton and Wheeler 1983), beneficial effects of extensive reading were treated descriptively, as accepted common knowledge, with no empirical evidence.

It was against this backdrop that I decided to try to measure what everybody was so sure of--if I manipulated one variable, extensive reading, could I measure significant English level gains over a four month period?

Instrumentation

A cloze test was chosen for its ease of construction and its high correlation with standardized English level tests (Aitken 1975; Stubbs and Tucker 1974; Oller 1973; Hanania and Shikhani 1986). A narrative passage from Jack London's To Build a Fire was chosen. The first three and the last three sentences were left intact. Fifty words were deleted at intervals of every seven words. The deletions included articles, pronouns, prepositions, adverbs, verbs, adjectives, nouns, and conjunctions. It was pretested on two native speakers who both scored 30 on exact response scoring.

Subjects

Four classes from the Department of Foreign Languages and Literature at Tunghai University, Taiwan, were selected to participate: day school sophomore, night school sophomore-experimental, night school sophomore-control, and day school freshman remedial. Only the two night school classes were paired for experimental purposes, however, since they differed only in the treatment.

The night school classes had nearly an identical makeup in terms of age, sex, prerequisite courses taken, and current course load. In addition, they used the same textbooks in each course, and followed the same syllabus.

The teacher of the control group was chosen because of her similarity to the teacher of the experimental group--both were non-authoritarian, outgoing professionals in their thirties who enjoyed excellent rapport with their students.

Experiment

Night school sophomore-experimental and night school sophomore-control were each taught two hours of deductive grammar, two hours of paragraph level composition, and two hours of functional conversation a week. Night school sophomore-experimental read (as did day school sophomore and day school freshman remedial) an average of 12 simplified readers over a 4-month semester, while night school sophomore-control did not.

The readers were mostly fiction, ranging from 600 headword to 4000 headword level. Students, except for day school freshman remedial, chose them based on interest and level. The only requirement was that the students must be able to go through them relatively quickly without the constant hindrance of a dictionary, and that the students must read for pleasure.

Results

All classes measured an increase measured along four criteria: acceptable response median (ar median), acceptable response mean (ar mean), exact response median (er median), and exact response mean (er mean). All three classes which utilized extensive reading made greater gains than the control group along the four measures, with one exception--night school sophomore-control and day school sophomore both registered +2 gain on exact response median. This was a greater percentage gain for the control group since it started at a lower raw median.

Comparing the control and experimental groups, raw and percentage gains made by the experimental group were substantially higher along all four measures than gains made by the control group. Experimental group outdistanced control group along the four raw measures (ar median, ar mean, er median, er mean) +100% +159% +100% +64%; experimental group outdistanced control group in percentage gains +90% +159% +120% +60%.

Experimental outgaining control along acceptable response and exact response was statistically significant ($p<.02$; $p<.08$).

Day School Freshman Remedial (N=10)

	ar median	ar mean	er median	er mean
pre	10.5	12.0	6.5	7.7
post	(+4)14.5(+38.2%)	(+3.1)15.1(+25.8%)	(+3.5)10.0(+53.8%)	(+2.6)10.3(+33.8%)

Night School Sophomore-Control (N=16)

	ar median	ar mean	er median	er mean
pre	17.0	17.0	11.0	11.2
post	(+1.5)18.5(+8.8%)	(+1.7)18.7(+10%)	(+2)13.0(+18.2%)	(+1.4)12.6(+12.5%)

Night School Sophomore-Experimental (N==17)

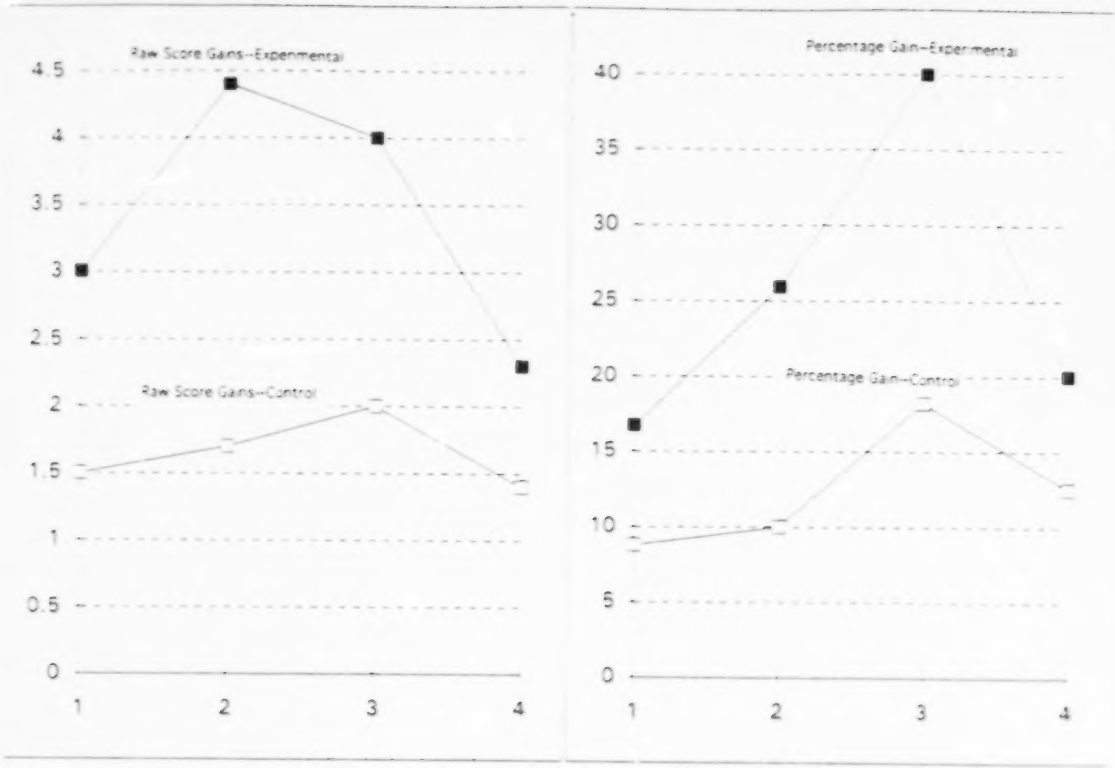
	ar median	ar mean	er median	er mean
pre	18.0	17.0	10.0	11.5
post	(+3)21.0(+16.7%)	(+4.4)21.4(+25.9%)	(+4)14.0(+40%)	(+2.3)13.8(+20%)

Day School Sophomore (N=14)

	ar median	ar mean	er median	er mean
pre	24.0	23.4	15.0	14.6
post	(+3.5)27.5(+14.6%)	(+4.3)27.7(+18.4%)	(+2)17.0(+13.3%)	(+2.8)17.4(+19.2%)

ar = acceptable response p<.02
er= exact response p<.08

BEST COPY AVAILABLE



Raw and Percentage Gains--1986 Study on the Effect of Extensive Reading on Language Acquisition, as Measured by Pre and Post Cloze Evaluation, Tunghai University, 1986

BEST COPY AVAILABLE

Conclusion

There is strong evidence that extensive reading promotes substantial language level increase within a short period of time as measured by cloze.

A technique that is effective obviously has many applications. Here is an activity that is not only student centered, but an activity a student can pursue independently and be relatively sure of positive results--that would make it not only effective, but cheap and convenient as well. Also, it is an activity that supplies teachers with an effective weapon, a trump card to use when confronted with stagnant, ineffective programs which, unfortunately, due to the moneymaking potential of English language institutes, abound worldwide.

General support for and agreement about extensive reading, backed by quantitative evidence, also begs a question: Why isn't extensive reading both encouraged and used in more EFL-ESL programs throughout the world?

Background--the history of the studies

I first became interested in extensive reading in 1984 while in a graduate TESOL program at the State University of New York at Albany. The Chinese students in the program possessed a communicative competence which appeared to surpass that of their counterparts from other countries. These students were primarily trained through audio-lingualism and extensive reading. The Chinese students were the only ones who had a grounding in extensive reading. Could extensive reading have made a dramatic impact on their communicative competence?

A subsequent controlled, empirical experiment offered statistical evidence to support a correlation between extensive reading and language acquisition (see Reading for Pleasure and Language Acquisition, Stephen Schackne, 1986). The subject of extensive reading has taken on renewed interest, spurred on, to a great extent, by Stephen Krashen's The Power of Reading. In this research review, Krashen cites evidence for a link between what he calls free voluntary reading and overall language competence. The renewed interest in reading and literacy issues has generated a cooperative replication study of Schackne's 1986 experiment.

At the University of Macau, a replication study using the same construct and measurement instrument as the 1986 study was set up. At Tunghai University a replication study using a much larger sample, and a more extensive battery of measurement instruments, was simultaneously launched.

Background--the literature

Literature exists, going back fifty years, to support the salubrious effects extensive reading has on language development. Specific studies, however, studying the effects of "pleasure" reading on second language development in EFL classes are rare. (I use the term "EFL," teaching English to second language learners in an area other than North America, the United Kingdom, or Australia, as opposed to "ESL," teaching English to second language learners in North America, the United Kingdom, or Australia. An English speaking environment and the variety of experiences students encounter in that environment would constitute too great a threat to the validity of the study).

In the 1986 study, I cited Jiang 1984; and Hubbard, Jones, Thornton and Wheeler 1983 as two articles dealing with extensive reading. However, I lamented the fact that in these articles (and others) beneficial effects of extensive reading were treated descriptively, as accepted common knowledge, with no empirical evidence.

Krashen has collected an extensive corpus, some of it empirical, dating back to 1948, citing correlations between free voluntary reading (interchangeable, I feel, with extensive or pleasure reading) and language development. The Krashen research review, however, deals primarily with native language development, not second language development.

We know that language input is a major factor in native language development, but can we generalize when two major factors, age and language, are manipulated; that is, is the same input factor that influences children in native language development relevant to adults learning a second language? That is what the University of Macau and Tunghai studies are trying to ascertain.

Instrumentation

A cloze test was chosen for its ease of construction and its high correlation with standardized English level tests (Aitken 1975; Stubbs and Tucker 1974; Oller 1973; Hanania & Shikhani 1986; Hinofotis 1979). A narrative passage from Raymond Chandler's "A Small, Good Thing" was chosen. The first and last sentences were left intact. Fifty words were deleted at intervals of every seven words. The deletions included articles, pronouns, prepositions, adverbs, verbs, adjectives, nouns, and conjunctions.

Subjects

Two classes from the entering freshmen class at large at the University of Macau, Macau were selected to participate. The classes included students who will major in a variety of subjects, but who are required to take a freshman efl course. The control and experimental classes had nearly an identical makeup in terms of age, sex, and current course load since freshman students at the University of Macau have a common first year. Prerequisite courses might differ slightly since Macau high schools offer a Chinese and an English track. However, the English 110 course, which the subjects were taking, draws most of its students from traditional Chinese track programs. The same textbook was used for experimental and control group. The same teacher taught experimental and control group.

Experiment

Two freshman English 110 courses were taught two hours and forty minutes of efl a week. The course covered all four skills with a minor emphasis on paragraph writing. Interactions II--A Reading Skills Book was bought by all the students. The experimental class read an average of 11 simplified readers over a 14-week semester, while the control class did not.

The readers were Longman and Heinemann graded readers, mostly fiction with some biography and general interest, ranging from stage 1 to stage 6 (300 to 3000 headword) level. Students chose books based on interest and level that was accessible to them. The only requirement was that the students must be able to go through them relatively quickly without the constant hindrance of a dictionary and the students must read for pleasure.

Results

The control class registered raw and percentage gains along three criteria, acceptable response median (ar median), exact response median (er median), and exact response mean (er mean). The control group registered a slight raw and percentage loss along one criterion, acceptable response mean (ar mean).

The experimental class registered raw and percentage gains along three criteria, acceptable response median (ar median), acceptable response mean (ar mean), and exact response mean (er mean). The experimental group registered no raw or percentage gain along one criterion, exact response median (er median).

The experimental class made greater gains than the control class along three out of four criteria, acceptable response median (ar median), acceptable response means (ar means), and exact response mean (er mean). The control group made greater gains than the experimental group along one criterion, exact response median (er median). Subsequent T-testing determined that although the experimental group made a slightly greater average gain along the criterion of exact response, there was no significant statistical difference in experiemtnal-control group gains along the criterion of exact response. However, the experimental group significantly outdistanced the control group along the criterion of acceptable response, $p<.03$.

English 110--Control (N=11)

	ar median	ar mean	ar median	er mean
Pre	16.0	16.9	12.0	12.3
Post	(+1)17.0(+6.3%)	(-.2)16.7(-1.2%)	(+3)15.0(+25%)	(+1)13.3(+8.1%)

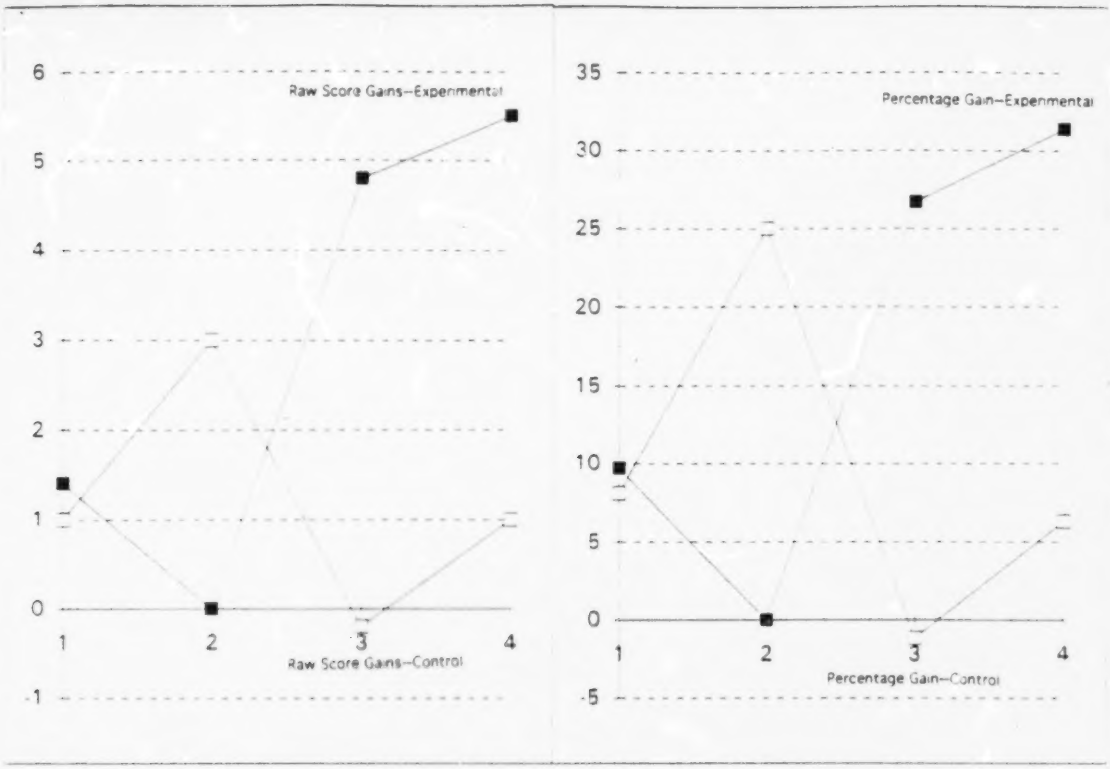
English 110--Experimental (N=10)

	ar median	ar mean	er median	er mean
Pre	17.5	18.0	15.0	14.5
Post	(+5.5)23.0(+31.4%)	(+4.8)22.8(26.7%)	(+0)15.0(+0%)	(+1.4)15.9(+9.7%)

ar=acceptable response
er=exact response

BEST COPY AVAILABLE

experimental group made statistically significant gains over control group in ar response $p<.03$.



Raw and Percentage Gains--1994 Study on the Effect of Extensive Reading on Language Acquisition, as Measured by Pre and Post Cloze Evaluation, University of Macau, 1994

Conclusions

There is evidence that extensive reading promotes language level increase within a short period of time as measured by cloze. But, even a cursory examination of both the 1986 and this study reveal that the 1994 results are not nearly as pronounced as the 1986 results.

Let me briefly address the factors that might (or might not) have skewed results. First, our original sample (control N=16, experimental N=12) pre tested at about the same level, the control group and the experimental group were within a point of each other along four criteria--median exact response, median acceptable response, mean exact response, mean acceptable response. Through drop-add, our final pre-post test sample was control N=11, experimental N=10, with pre test scores varying from one to three points along the four grading criteria, with the experimental group starting at a higher level along all four criteria. One could question the experimental group starting at a higher level may have a better foundation or aptitude for language development than the control group. Conversely, starting at a lower level, the control group may benefit from having a greater statistical range to improve, sort of the opposite of the "Hawthorne effect."

We must also add that the measurement period was two weeks less than in 1986, and the average number of books read was one less.

Thirdly, the statistically significant results occurred in the area of acceptable response, the scoring criteria, due to both grammatical and contextual factors, most susceptible to scoring error. One could question that scoring error accounted for the different results along the criteria of exact response and acceptable response, although, to date, no significant scoring errors have been discovered in the 21 scripts. Conversely, one could argue, that acceptable response gains on the part of the experimental group simply indicate superior creativity in productive language skills.

The research on the two scoring methods is still unclear--Oller (1972) and Hinofotis (1976) suggest that the acceptable-word scoring method yields more reliable scores and provides more accurate information about esl proficiency levels; however, Stubbs and Tucker (1974) and Oller et al. (1974) indicate very little difference between the two scoring methods. Brown (1978) felt that the acceptable-word method was more appropriate for measuring productive language skills, but there is no consensus for either method.

Final Word

Colleague Guo Yi Qin will continue the study through the spring semester, 1995. Results will be tabulated, but not formally presented, to see if experimental group keeps outdistancing control group.

Based on 1986 and 1994 results, a replication study(s) should be undertaken with the following suggestions:

a) sample size > 15 consistent from pre to post; possible multiple class study increasing sample size

b) pre testing any cloze instrument with native speakers to eliminate any linguistically controversial items

c) possible multi-instrument evaluation along different individual skill areas

d) sample taken from non-Sinitic language group in home country; e.g., Brazilians or Hungarians.

References (1986 Study)

Aitken, K.G. (1975). Problems in a cloze testing re-examined. TESOL Reporter, 8:2.

Hanania, E. & Shikhani, M. (1986). Interrelationships among three tests of language proficiency; standardized esl, cloze, and writing. TESOL Quarterly, 20, 97-109.

Hubbard, P., Jones, H., Thornton, B., & Wheeler, R. (1983). A Training Course for TEFL. Hong Kong: Oxford University Press.

Jiang, H.S. (1984). Teaching extensive reading. Forum. 22, 4, 37.

London, J. (1964). To build a fire. In Grindell, R.M., Marelli, L.R. & H. Nadler (eds.). American readings (192-193). New York: McGraw-Hill.

Oller, J.W. (1973). Cloze tests of second language proficiency and what they measure. Language Learning 23:105-118.

Stubbs, J.B. & Tucker, G.R. (1974). The cloze test as a measure of english proficiency. Modern Language Journal 58: 239-241.

Zukowski-Faust, J., Johnston, S., Atkinson, C., & Templin, E. (1982). In Context. New York: CBS College Publishing.

References (1994 Study)

Aitken, K.G. (1975). Problems in a cloze testing re-examined. TESL Reporter, 8:2.

Brown, J.D. (1978). Correlational study of four methods for scoring cloze tests. Master's thesis, University of California, Los Angeles.

Hanania, E. & Shikhani, M. (1986). Interrelationships among three tests of language proficiency: standardized esl,cloze, and writing. TESOL Quarterly, 20, 97-109.

Hinofotis, F.B. (1976). An investigation of the concurrent validity of cloze testing as a measure of overall proficiency in English as a second language. Doctoral dissertation, Southern Illinois University.

_____ (1980). Cloze testing: an overview. CATESOL Occ. Papers, 6:51-55.

Hubbard, P., Jones, H., Thornton, B., & Wheeler, R. (1983). A Training Course for TEFL. Hong Kong: Oxford University Press.

Jiang, H.S. (1984). Teaching extensive reading. Forum, 22, 4, 37.

Krashen, Stephen. (1994) The Power of Reading. Englewood, Colo: Libraries Unlimited.

Oller, J.W. Jr. (1972). Scoring methods and difficulty levels for cloze tests of proficiency in English as a second language. Modern Language Journal 56, 3, 151-157.

_____(1973). Cloze tests of second language proficiency and what they measure. Language Learning 23, 105-118.

_____, Irvine, P.,and Atai, P. (1974). Cloze, dictation, and the test of English as a foreign language. Language Learning 24, 2, 245-252.

Schackne, Stephen. (1986). Reading for pleasure and language acquisition. Unpublished research. Tunghai University.

Stubbs, J.B. & Tucker, G.R. (1974). The cloze test as a measure of English proficiency. Modern Language Journal 58, 239-241.

READERS

Treasure Island	R.L. Stevenson	Longman Structural	Stage 3
The Picture of Dorian Gray	Oscar Wilde	Heinemann	Elementary
Computers	Lewis Jones	Longman Structural	Stage 4
The Face on the Screen and Other Short Stories	Paul Victor	Longman Structural	Stage 2
L.A. Detective	Philip Prowse	Heinemann	Starter
The Great Gatsby	F. Scott Fitzgerald	Heinemann	Intermediate
Brave New World	Aldous Huxley	Longman Structural	Stage 6
Thunderball	Ian Fleming	Longman Structural	Stage 5
The Mystery of the Loch Ness Monster	Leslie Dunkling	Longman Structural	Stage 1
David Copperfield	Charles Dickens	Longman Structural	Stage 3
Space Invaders	Geoffrey Matthews	Heinemann	Intermediate
Go Between	L.P. Hartley	Longman Structural	Stage 6
Grapes of Wrath	John Steinbeck	Heinemann	Upper
Round the World in Eighty Days	Jules Verne	Longman Structural	Stage 3
Operation Mastermind	L.G. Alexander	Longman Structural	Stage 3
8 Ghost Stories	S.H. Burton	Longman Structural	Stage 3
Cider With Rosie	Laurie Lee	Longman Structural	Stage 6
A Scandal in Bohemia	Arthur Conan Doyle	Longman Structural	Stage 4
Great Expectations	Charles Dickens	Heinemann	Upper
Ghandi	Donn Byrne	Longman Structural	Stage 3
Cry the Beloved Country	Alan Paton	Longman Bridge	

Modern Short Stories	G.C. Thornley (ed.)	Longman Structural	Stage 6
On the Beach	Nevil Shute	Longman Structural	Stage 5
Elvis Presley	Jeremy Harmer	Longman Structural	Stage 1
Pele	Noel Machin	Longman Structural	Stage 1
The Thirty Nine Steps	John Buchan	Longman Structural	Stage 4
The Prisoner of Zenda	Anthony Hope	Longman Structural	Stage 4
Desiree, Wife of Marshall Bernadotte	(no other information available)		
Munich Connection	(no other information available)		

BEST COPY AVAILABLE